ME3023 Lecture - Chapter 4

Probability and Statistics

Theory and Design for Mechanical Measurements

5th ed. by Richard Figliola and Donald Beasley

Tristan Hill - Tennessee Technological University - Fall 2020

- ## 4.1 - Introduction

  Who has taken a statistics class before? In college? In high-school?

  What does probability and statistics have to do with mechanical engineering?

  For a given set of measurements we want to quantify...

    – a representative value that characterizes the average of the measured data set

    – a representative value that provides a measure of the variation in the data set

    – how well the average of the measured data set represents the average of the entire population

- **4.2 - Statistical Measurement Theory**

    - Where does the measured data set come from?

    - Sampling refers to repeated measurements of the measured variable under fixed operating conditions.

    - We will ignore systematic error for this discussion, is this valid?

    - Instead we will focus on random error, its affects and how to quantify it.

    - Question: If the error is really random error, what is the average error?

– We want to estimate the true mean, x' from repeated measurement of x.

– The true mean, x' is the average of all possible values of x. We never actually get this!

– Through sampling we can find $\bar{x}$, the sample mean value of x. We do get this!

– As our sample size increases, $\bar{x}$ approaches x'.

$$x' = \bar{x} \pm u_{\bar{x}}$$

– Therefore, the sample mean $\bar{x}$ is the most probable estimate of the true mean x'.

– $\pm u_{\bar{x}}$ is the uncertainty interval in that estimate at some probability level, P%.

– The uncertainty interval is the range about $\bar{x}$ that you would expect x' to lie.

- # Histogram

  "A histogram is an accurate representation of the distribution of numerical data. It is an estimate of the probability distribution of a continuous variable (CORAL) and was first introduced by Karl Pearson.[1] It differs from a bar graph, in the sense that a bar graph relates two variables, but a histogram relates only one. " - Wikipedia

- # Probability Density Functions

  "... a probability density function (PDF), or density of a continuous random variable, is a function whose value at any given sample (or point) in the sample space (the set of possible values taken by the random variable) can be interpreted as providing a relative likelihood that the value of the random variable would equal that sample ... the PDF is used to specify the probability of the random variable falling within a particular range of values, as opposed to taking on any one value. This probability is given by the integral of this variable's PDF over that range—that is, it is given by the area under the density function but above the horizontal axis and between the lowest and greatest values of the range ..." - wikipedia

  - The frequency with which the measured variable assumes a particular value or interval of values is described by its probability density function.

  - If a central tendency exists we should be able to see this in the probability density function.

  - As binsize of the histogram of the data set goes to zero this becomes the probability density function.

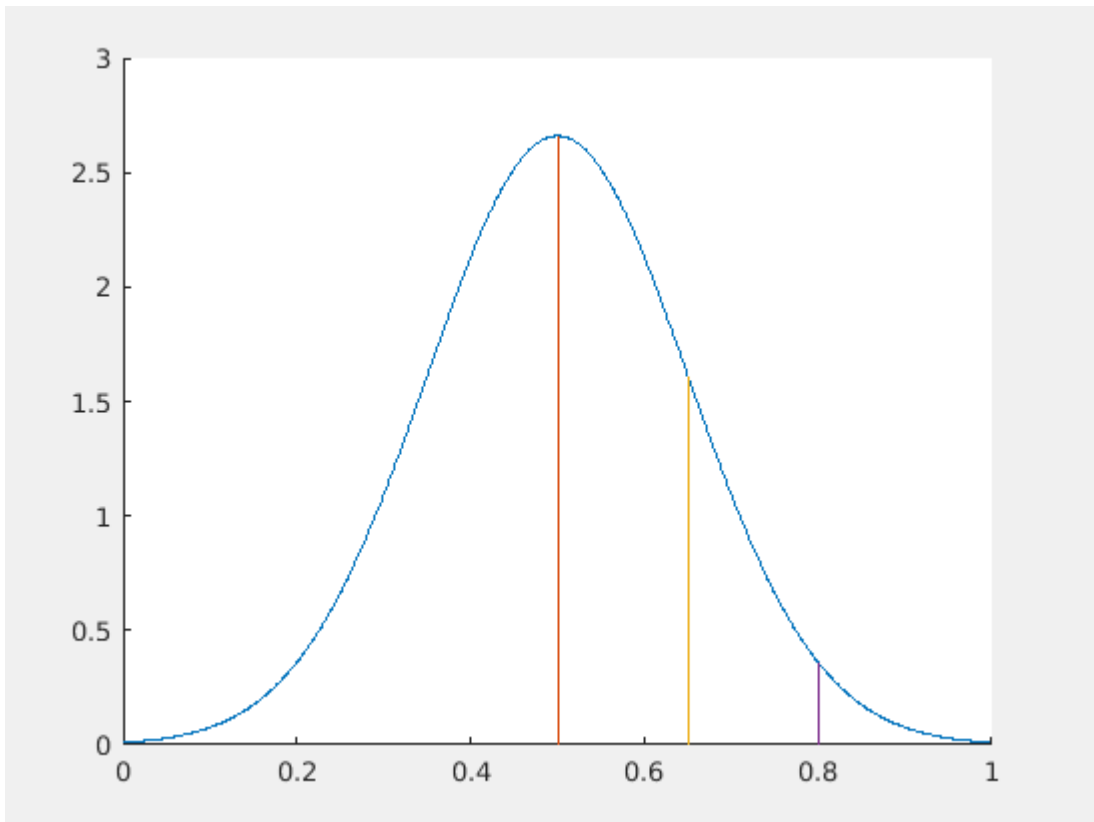- **4.2 - Describing the Behavior of a Population**

The true variance is:

$$\sigma^2 = \int_{-\infty}^{\infty} (x - x')^2 p(x) dx$$

For discrete data this becomes:

$$\sigma^2 = \lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} (x_i - x')^2$$

The square root of the **variance** is the **standard deviation**.

$$\sigma = \sqrt{\sigma^2}$$

- Now let's do an example.